

February 2010

Brewster Kahle

Internet Archive – Co-Founder & Director

Interview by: Daniel Erasmus, Managing Director, The DTN

I'm Brewster Kahle, founder and digital librarian of the Internet Archive, and have been involved in a lot of the technologies and companies that have created the Internet that it is today. Started my career after MIT at Thinking Machines, building big computers so that we could actually do something with them, then made the first search engine that was a broad-based search engine, then made that available on the Internet as a distributed publishing platform called WAIS. I started that as a company, sold that to AOL. Then, once we had publishing starting to move on the Internet, then we could start to catalogue and collect all of the information that is the Internet, that is available on the Internet, to build a great library. Built Alexa Internet, which catalogues all websites, and the Internet Archive, which is the idea of a precursor to a Great Library of Alexandria, version 2.0.

So we're still in the midst of this. This is 2010. We're still in the transition of putting everything online, not just this organisation but the whole world, and starting to put it together in such a way that anybody can have access to anything at any time forever. And so, all published information is starting to be available to people at any time, any place. And that's a very exciting place to be, and worthy of sort of a generation's efforts, I would say.

01:34 INTERNET BUSINESS MODELS

You can take the business models that have sort of evolved over how do you get information out there? And there's advertising. There's royalty-based systems like books, and there's subscription-based systems, sort of who you know, a little club to go in and to be a part of. I like the royalty-based systems because it means that there's a simple buying structure. Anybody can buy anything, and there's a very

easy method of distributing the proceeds of that. Advertising-based industries tend to consolidate, so you get things like newspapers, magazines, television. Those are sort of the advertising-based. And now you get these search engines that suck the air out of the room for other people because they can control the ad-selling network. And it's always been that way, so you end up with something quite different than you thought you were going to get, very consolidated, centralized power.

2:43 THE IMPORTANCE OF DISTRIBUTING POWER

In the information age – so, we're not talking about just railroads or software, where actually it is important to have distributed ownership of transportation and our software infrastructure. And so, we wrestled with the monopoly structures in those areas before. But in the information area, it's even more important that if somebody really is able to control what people think, how people are educated, then I'm not sure you can get your way back out of it, that if you have just a few players being able to control things at that kind of level, such that how thoughts evolve is influenced, even if it's not diabolically controlled by one or two corporations, it will influence the evolution of our society and who we are as people in a very corporatist, centralised model.

3:47 WHO OWNS THE INTERNET

If we're trying to figure out sort of where is the ICT infrastructure information technology going in 10 years, sometimes it helps to go back 10 years and try to see where we came from. Let me pose a question. Who owns the Internet? Who owns the Internet? And I

think our answer to that question – both our perceived answer and maybe what's real, and how that might be changing, our perceptions might be changing – could give a really good mirror onto how we see what's going to happen. And how we see things going to happen might actually influence what's going on.

So, 10 years ago, the answer to, “Who owns the Internet?” was laughed at. There was no one that owned the Internet. It was always a very – it's a scripted answer. It was like it's a network of networks. It's an open structure that anybody can go and create things on. You didn't need permission to go and build things on the Internet, and we had fantastic things to be created in that world, whether it's e-mail, the World Wide Web. The peer-to-peer networks in '99, 2000 were coming out so that you had the greatest music library ever made, which was built on top of Napster. People just poured their time into going, putting, making all of this available. It was the sort of the fundamental premise of the Internet.

So, now we're 10 years later and trying to predict what's going to happen in the next 10 years. Let's take that question again, and I think it might be illuminating, and both as a perceived, “Who owns the Internet?” Is that question even more relevant now than before? I would say yes, that people are starting to answer that question that there is someone that owns the Internet, or some piece of the Internet, or some of it's controlled. Maybe it is, maybe it isn't, but people are starting to think that way. But let's take it as no, people don't own the Internet, and new and wonderful things are coming up because it's open, because there's a mechanism of getting to a distribution channel. You can get your ideas out there.

I look at the top 10 websites, and I happen to know them from year 2000, and I'm pretty sure what they are now. In the year 2000, the top 10 websites were 20 percent of the attention on the Internet – the clicks. If you count all the clicks of all the people of the world, 20 percent of them go to 10 organizations – pretty astonishing. That's a real concentration of power. They were mostly things like search engines. AOL was in there. I guess sort of the normal cast of characters that you'd kind of imagine – eBay was in that list. Amazon was in that list in the year 2000.

Now, we have some of the same characters, but we have Facebook in there. We've got Wikipedia. We've got a search engine from China. So, there's some shifts that are going on in the top 10 websites, but what also has changed is the amount of time and attention people put towards the Internet. Television has continued to fall away as an influential medium, so people are spending more and more time on the Internet. So, while there's still – there are these important organisations in there, but they do change. They're not controlled by a few organisations that are going to stay in place forever, that the open nature of the Internet and ICT technology has made it so that there is upward mobility, even into the top ranks.

So, we have these worlds. Let's take the answer as somebody does control the Internet, and then somebody doesn't control the Internet. I think this debate in 10 years, having a question, will be very important again, but even our perception in the next 10 years. If you answer that somebody does own the Internet, what does that mean to our organisations, our role in society and the like? And I think that that, companies becoming either defensive or open will start to really pivot on this idea, is there a man? Is there somebody that you're going to have to pay attention to or please to be able to be successful, or are

we really in an area that a new idea can come about and have their place?

8:39 FROM TECHNOLOGY TO CONTENT

Another thing that's changed in the last 10 years is we've shifted from a technology focus to a content focus. We used to be that really it's sort of who organised server farms and the like. It was who controlled the protocols, instant messaging 10 years ago, were the sort of the emerging technologies. Now, we've got more and more people going and figuring out how to own other people's information, whether it's YouTube or Facebook, where there are whole environments for existing.

Now, we're not avatars running around in a big, 3-D world. That's happening, but it's not the dominant force yet. It might be in 10 years, but we're starting to see whole environments that are built in platforms for people to interact in, but more importantly, put critical information about themselves. Data mining is all the buzz in the last 10 years within the government sectors and corporate sectors, but when we really look at it, very few people, very few organisations know how to use the information that is gushing out of our web servers, out of our IT technology. And those organisations may become more and more important, and more and more consolidated.

More of the publishing infrastructure, what we think of as publishing – it might have been old newspapers or magazines or books in the old day – it might be something quite different that involves very large tracts of data, and organisations that control and know how to exploit these classes of information.

So, what does it mean to have a social network where people know who all your friends are? What good is it? Who cares? What does it mean to be an organisation, to have something proprietary in a world where your organisation's information is going through somebody else's e-mail system? What does it mean when you're not using Excel and Word so much anymore as you're using these hosted services, where who knows quite where your business plans are going? And who's able to exploit that in interesting and new ways?

Even if it's the threat of exploiting it, can help play into people's hopes and fears as to what it is our IT infrastructure is doing for us, or are we really just feeling like we're kind of a cog? And just the idea of freedom of open opportunity is an incredibly important myth towards creating an interesting world that's served us very, very well. We've had some really interesting technologies built. I think we have an opportunity to do much more of it, but I think it's going to be people's answer to the question, "Who owns the Internet?" that really will help divide the world.

11:41 A TRUSTING WORLD

An amazing shift that's going on among all of us that are in our 40's is what our children are doing on the Internet, and how much trust there is in what's going on. And what they see is the benefit of going and having their social experiences be completely exposed. Their photos, who they're friends with, what they're doing all go through other people's computers. Those of us that were brought up by parents that had gone through World War II, the Depression, the McCarthy era of you can get in trouble, based on your associations – we were warned about these sorts of things, and it was just part of the air that we were brought up in.

But this next generation is finding much more benefit to just sharing it all, to going and having your photos on Flickr, your Facebook. And there is just a – there's a trustingness that's going on at the populist level that may be warranted and may be not, but it's certainly going on. And there are organisations that are starting to learn how to exploit this trust that is built up, in ways that not right now seem like targeting of advertising. There's certain interactions with the government, where there's more and more information being shared with the government's intelligence and law enforcement.

Since I run a digital library, we archive the World Wide Web, so people have – we have probably the only most public copy of the web that's existed over a decade's time. We collect a snapshot of the Web every two months, and it's used by, oh, about 500,000 people a day come in and use this resource for all sorts of things. And we get an increasing number of lawyers coming and demanding information from us, and it's kind of interesting because it's all publicly available. They could just use it, but they come and do demands with subpoenas and thuggish stuff that we have to go and do something for them.

So, why is this interesting? One, it's just the growth curve, just the number of organisations that are using these resources to sue each other, that we're seeing this. And since we're a public entity, I have no problem going and telling you this. This must be going on with Facebook or Gmail or Yahoo! and whole teams of people are just dealing with law enforcement all of the time. Amazon will tell law enforcement, with a proper court order, what books you've read.

So, those sorts of things that strike fear in us, that we really know in the 20th century is absolutely happening now. And we have a world that's very trusting. The good thing about this is all sorts of great things

happen. You find all sorts of interesting people are a member of many different communities that overlap each other geographically, that you have several facets of yourself, and it's OK to be developing those at the same time.

Our work environments are becoming much more fluid between what's your personal and work environment, and we're leveraging all of that so that – somebody yesterday was helping in sound design in our new building, but he's a programmer doing great things on cluster development, but he happens to also do a lot of things in clubs at night. So, he knows a lot about this and he's able to leverage his hobbies and his other interests that are highly developed, in large part because of the Internet and the other people that are able to interact in his work world. It makes him a much better employee, and it makes us much more interesting people because our information sources are so wide and diverse.

It's a trusting environment, which is very, very good, but might really come back around in some interesting way. These next 10 years are going to be interesting in this way. How do corporations deal with it? How do governments deal with it, when there's an export of their information and is some level of control? I live in Silicon Valley. It all seems pretty good to me that the music industry really looks like it's being majorly contorted by Apple Computer. Google has got information on people from all over the world. We have Yahoo! here. We have Facebook. It all seems kind of good, but when you're at the centre. How does this feel if you're not feeling like you're at the centre? What if it feels like I don't really trust these guys to help out? How is it going to evolve? Is it going to be more centralised, less centralised? Is it going to come back to bite us?

And it's at the content level that I think is important, not so much at the technological level. Yes, it's important who owns and controls the transport systems, but who controls the filters? What are the new applications that are coming out? Does everybody – is there going to be a new way of communicating that's able to be built on iPhones, or is Apple going to control that, such that it's not going to be kind of what they want to have out there?

The Macintosh and the PC allowed anybody to write a piece of software and distribute it for free. It was hard to make money on it in the early '90s, but you could get it in front of millions of people. That used to be very difficult to do through the retail chains and controlled distribution structure, but you could get this thing out there.

In these new worlds that are really dominated by things like video games, where it's really controlled environment – phones, really controlled environments, there's some hand wave about, “Oh, well, there's this network we have to protect.” Is that really the issue? Are we in an era where new applications, new companies, new technologies for managing social interactions are going to be able to blossom? Are we going to build new and different navigation structures for taking this vast wealth of information, of personal, corporate and wide area of information that's becoming available, and making new answer engines that go and learn from these and extracts facts, finds contradictions? Those are some of the new things that are coming out.

The Wolfram Alpha engine that is really what the AI community 30 years ago was dreaming of, is now starting to come about. The reason why they're coming about now is the information is available, that you can go and just have curators that go and pull information from a wide

variety of sources to make new information sources. 10 years from now, are there going to be more of these or less of these? Our answer of what's our perception of that will separate us. But also, the reality of it, I think, is our opportunity to shape our future.

And I do believe that it's – the open thing just continues to surprise us, to shock us, that it's the kid in Lithuania that comes up with something that seems a little risqué, a little bit odd, but is what is going to blow things up and make it a very interesting 10 years. How do we position ourselves with that kind of change? The same way we've looked back – how were we doing it the last 10 years? Did we do a good job? If not, maybe look at your kids and how trusting they are.

19:19 SILICON VALLEY AND THE WORLD

So, some of the world changes that are going on, the role that Asia is increasingly playing, Europe, Middle East, how is it evolving? I get a bit of a point of view from where I sit by running a digital library and working with basically information flows. Where is it coming from? Who's doing what? Who's investing in the content layer of the Internet, not so much the technology layer, but the content layer? And where are ideas coming from? And increasingly China is where at least the hardware technology is really shifting, so that if you really want to know what the future of tablets look like, go hang out in Taiwan. Everything's so dependent on them. They're really setting the pace. Yes, maybe it still has brand names that still look American, but it's over. It's shifted, and when the hardware manufacturing, then the design follows and it really wraps around, what it is we can do next. And you get changes such as the touch screens that will really radically change what's going on, but it will be dominated by the Asian organisations.

And I think that's great. Let's spread it out somewhat. The Silicon Valley just had a major win in the last two years. Wikipedia moved to San Francisco – interesting. So, it used to be in Florida – great idea blossomed – but they felt kind of out in the woods. So, the non-profit content organisations still are finding their way to San Francisco, and I think it's because that there's a sort of balance between openness and proprietary control that we still see explored in the Bay Area.

The Grateful Dead gave away their concert recordings and did very well that way. Would that have happened in Los Angeles? Would that have happened in London? I don't think so. It happened in San Francisco 30, 40 years ago. The World Wide Web, CERN, Tim Berners-Lee went to Boston to go and do the World Wide Web Consortium in Boston, but it was really Silicon Valley that really took it and ran with it. But the Electronic Frontier Foundation, Internet Archive, Wikipedia, the Mozilla Foundation – these organisations that are building infrastructure, new operating systems – Linux Foundation – these things are in the Bay Area.

So, does it mean that all the work is going on here? Absolutely not. More and more of these communities are worldwide, and more and more of our organisations have got interacting communities that are everywhere. But there's still something about the Bay area's ability to do this balance. Is it shifting to other places? Of course, all eyes are on China, Taiwan, Japan. How that will shape and influence, I don't know, but I think more and more of us are going to start learning Mandarin. More and more of us are going to spend time there. Where we might have spent time in Paris, we're going to put on our schedules going to Hong Guo [SP] or going to Beijing, and just to hang out, understand, go to conferences, learn, absorb. What does the world look like from over there?

And certainly I've spent a couple months this last year, just spending time over there. And I've spent time over in Egypt and in Europe as just we have to, to just broaden up, to get an idea of the influences that are going on, because it's influencing our culture, it's our technology, the content, and the point of view on lives is all coming from these other places, as well now.

23:40 DIGITAL DARK AGE

So, we're prehistory in the sense that there's going to be a gap in our knowledge, because the writers are writing on word processors that you throw out. Do you have your electronic mail from 10 years ago? Probably not. It's too hard to keep these things now, but what does that – that might mean there's a real gap in the record, or it's a very selective, weird record. It's very deep in some ways and not in others. So, if it's in other people's corporate memory banks, it might be there by now.

One of the great early websites of homesteaders, the Facebook of the day, was called Geocities, and it shut down. Yahoo! shut it down. They bought it for about \$400 million back in the dot-com days, and last year they shut it down. And they, to their credit, they let us know at the Internet Archive that they were going to shut it down in nine months, and so they said, "Crawl like nuts." And we said, "Can you give us a list of all of the homesteads on it?" And they said no, because that would sort of violate what they thought the privacy interaction with their user was at the time. But if we knew of the site, if we knew of it from any which way, if there was any links into it, then feel free to crawl, to be able to make sure that there's original content set.

And that's not just sort of historically interesting. It's documents that are referenced all over the web all the time. The Internet is vastly

trusting of other people's sites. Websites have so many external links that we're all dependent on each other to survive. And when a big hunk of the web like Geocities goes away all at once, we have troubles, in terms of how the Internet's built. So, this is what some people are thinking of as a digital dark age.

25:49 THE INTERNET ARCHIVE

The Internet Archive is a digital library in the modern age. We're becoming a little less odd to people – I don't know, maybe as we're getting older. It's 13 years old. Been archiving the World Wide Web with a snapshot every two months for the last almost 15 years. And this, we've made into the Wayback Machine, so that people can go and use a URL to find old websites. It's the out-of-print web. Statistically, we have 150 billion web pages from 50 to 100 million websites, that the average life of a web page is about 100 days before it's either changed or it goes away. So, it's a really evolving mush out there. And being able to have access to the older material means that you actually have a reliable web. And it's used by, as I said, about a half a million people a day.

We also started collecting television, so as much as it sort of seems like yesterday's medium, it's still very important, and people had not been doing a very good job archiving. So, we've been archiving 20 channels 24 hours a day – Russian, Chinese, Japanese, Iraqi, Al Jazeera, BBC, CNN, ABC, NBC, CBS. And it's sitting dark. It's not accessible, mostly from a money perspective. We're trying to get it available.

We also collect movies and music, and people often upload this, or we digitise them and make them available permanently as reference collections. And we're also archiving books. We've got – we're scanning about 1,000 books a day in 20 scanning centres, including a

scanning centre at the Library of Congress, one in London, Scotland – 20 of these things in five countries to go and make it so that books are available again.

So, mostly we have now 1.8, 1,800,000 out-of-copyright books, so anybody can download those. And we're also scanning things that are in copyright and making those increasingly available to the blind, because there's copyright exemption for that, and we will be launching a lending library so that kind of like when we were growing up, you could go and get materials from the library and then bring them back. We're focusing on the out-of-print, so that the in print, the publishers have sort of room to try to build a business model of books on the Internet.

So, the idea of universal access to all knowledge is within our grasp, and we're making progress on that, as well as some other organisations. Google very dominantly is working towards this, but that idea of building the Library of Alexandria Version 2.0 is no longer just a hand wave. Significant percentages of all the published materials ever are now in digital form, and in different ways, are available to people.

29:01 DESIGN PRINCIPLES FOR THE DIGITAL WORLD

So, as we're making huge collections of information online, we're starting to see an evolution of a couple different approaches to it. There's one that comes from a tradition of libraries and one that comes from a tradition of products. And neither one of these are quite right. Libraries in the past didn't have information like social network information in them, and corporations had products that they designed, built, and owned, and now they're starting to accumulate other people's information into them.

And these two different points of view are coming out with very different futures. My favourite, we're very much involved in right now, and the big tussle is around books. As books are going online, Google is digitising these books, and they're doing it very fast in such a way that they can put it into their search engine and exploit it and sell it. And they're really taking sort of a top-down, "We own it all" control. Everything is contractually regulated. All of the user's trails are audited and kept, so that what everybody reads is owned by a corporation, the terms of service.

So, that's one approach, which is really coming from a product design background. It just expands to be the library of all books. Then, there's, you take another approach which is the Open Content Alliance, the Internet Archive. We're working with about 150 libraries to digitise books from all over the world. We're digitising about 1,000 books a day, ourselves. We're working with projects in India and China. It's the open world. It's a very open, exchange-oriented, a lot less top-town control. We don't keep usage information. We don't know who read what. We know how many people read something, but we don't know who they are. We never kept it because it's part of the tradition of coming from a library, and it gets spooky when you can go and track who reads exactly what.

The ownership is much less dictatorial. We share things a lot more. We'll push a million books back and forth from continent to continent just to make sure there are backups in other organisations, because we see longevity as part of us. So, corporate control will live forever – trust us – versus the distributed, open world. And so, as books are going online, we're seeing a competition of these two approaches. Which will win? I don't know. Will it be a hybrid? I doubt it. I think it

will really be dominantly one way or the other, and it's playing out now in a fairly dramatic way.

Music, when it went online, really went to one corporation: Apple. And yes, there's third-party sellers that are trying to play a catch-up role, but it's really owned and controlled at this point, the distribution of commercial music. And the only way around that is a bottom-up reinventing music completely. So, it never went to a distributed World Wide Web of music. It went to a centralised store of music.

I'd like to see a distributed world of books, a World Wide Web of books, a World Wide Web of information, educational materials. But the temptation of a single organisation, the opportunity of a single organisation, taking and controlling whole media types is one that can shape the next 10 years in a very interesting and important way.

32:45 CLOUD COMPUTING

Cloud computing is an evolution of this networked world, which is fantastically efficient. The idea of software as service, or as outsourcing of different parts of organisations, or allowing interactions, is causing people to be very excited. But pulling it apart a little bit does make sense. So, software as service – it's a lot easier to make an application that runs on a server than it is to try to make an application that is a compiled piece of code that distributes and runs in everybody's Windows machines or Macintoshes or Linux boxes. You can rev it much more quickly. You can put new ideas into a service so that if you were running a company, trying to put ideas out there, putting up a website – even a sophisticated one that has software that downloads part of your browser – is so much more efficient. And that's sort of thought of as cloud computing – software as service.

There's also other things called cloud computing, which is like the Amazon service of being able to be an outsourced IT system, and most of the start-ups these days are using those services so that they don't have to manage hardware. Hardware is really hard to manage, and it's hard to know how much you should get. It sort of seems all very old-world that you have to go and pre-invest in hundreds of thousands, millions of dollars worth of hardware infrastructure, when you could just use as much as you want to, this utility model.

The downside of that is well, it's on somebody else's server, and what does that exactly mean? But it's starting to be even more than just a software service or outsourced IT. There's starting to be these cloud – people are using these cloud services, meaning that there's these collaborative environments, shared information environments that are hosted, whether it's Google Docs, Facebook, Flickr. You're putting things in the cloud, and it's not just putting in the cloud because it's more efficient to go and store your things on somebody else's IT infrastructure. It's because there's new and added features that you get by having crowd sourcing of information connectivity, that you're putting your stuff with everybody else's stuff out there.

But the metaphor of a cloud, I think, is deeply deceiving, the idea that it's sort of everywhere and it belongs to everybody is absolutely not true. It actually does belong to somebody, and it is going to be exploited within their own mental model. And we haven't seen the rise of non-commercial clouds yet, or things that are a little bit more trusted, things that don't give you the heebie jeebies if you really think about what information they have in them, and what level of influence they could have by having this information.

So, clouds are starting to mean way too much, but I'm seeing it as these different areas, and the most powerful one is the interactivity between people that's enabled by having our software and services and our information in these combined mush out there, that then can be data-mined for good and for inappropriate reasons.

36:16 GOVERNMENTS HAVE TAKEN A BACK SEAT

I was involved in the earliest cluster computing search engines back before there was the Internet, helping the Internet come about. So, I've been involved in a lot of these things and seen a bit of perspective. What's surprising to me is how slowly things have changed. I think that really shows my failure more than – just how long it takes. The idea that we're in 2010 now and we don't have textbooks online is just, how could we possibly be here? What does this mean that the maturity of the technology is still so lame? That there isn't at least some base level of free Internet access to all wireless devices now in our cities is, I find, really hard to believe.

And it's happening, but so slowly and in such fits and starts and in clunky ways, that the government has really taken a backseat role where 20 years ago, a lot of this was happening in universities and national labs. A lot of it now is happening in either corporations or independent non-profits. What happened to the government? I mean, what happened to the next generation of information superhighway, and how do you level the playing field and sort of keep things under control?

The government is just absolutely checked out in every – in Europe and in the United States, there's not an investment in large-scale, interesting infrastructure projects in the universities. They're just

playing around the edges of how do you build services on top of Google in some putzy way? The new generation of supercomputers aren't coming out of these places. The national labs are trying to rent them from the Googles and Yahoos! and Microsofts out there. What's this about?

So, the things that I find surprising is the shift and roles between dot-edu, dot-gov, dot-com, dot-org. A thing that I found very surprising and very encouraging is the rise of the dot-org, the non-profit, that that's been the – it's nongovernmental, international organisations that are building infrastructure. That they operate on a set of principles of privacy, of transparency, that they are – people feel it's OK to go and invest their time and effort, whether it's hobby hours, to build a Wikipedia, or going and building Amazon's reviews. You want to be paid for that, so there's this new sector that I find very interesting, very surprising, and I'm very intrigued with where that's going to go.

I think we may be building another generation of government, if you will. They have their own rules and regulations. It's in its infancy, but these non-profits are interacting with for-profit corporations in very symbiotic, positive ways that this is, I think, a very interesting evolution. And I think it's come about because we've had basically governments that have just been off doing dumb stuff for the last 15 years – in the United States, the last 30 years. It's been in relative collapse.

40:02 HIGH TECH NON-PROFIT ORGANISATIONS

So, the high-tech non-profits are starting to really take off. They started really with some of the open source software moment. So, Richard

Stallman – the grandfather of this whole idea – he made his Free Software Foundation be a non-profit. Linux Foundation is, again, non-profit. So, in terms of those that are governance over open source software, they're in the non-profit sphere. There was an attempt with Red Hat, which is a commercial company, to be a commercial owner of Linux, and it failed. It failed. Java is within Sun and we'll see how that evolves. It's sort of an uncomfortable place.

But these non-profits are doing very well in these open source software areas, and they're starting to build full on operating systems, or browsers like the Mozilla Foundation is a non-profit building a very viable contender to browsers. There were browser wars 10, 15 years ago between big commercial players pounding it out. Now, it's open source is really the contender. What does this mean? Big hunks of what Apple ships is open source, and so, where Apple is leveraging it, it's owned and controlled, if you will, governed by non-profits.

We're seeing this now happen in the commercial sphere. Wikipedia, Internet Archive, the Public Library of Science – all of these are non-profits, and they come with rules. You can't pay people too much in them. They're not buyable, which is very important for longevity. I met with Mitch Kapor. Mitch Kapor started Lotus, the important – the spreadsheet – important guy. But I saw him at the 20th anniversary of EFF. So, he was one of the founders of EFF and the founders of Lotus, and I asked him, "So, what else have you been involved in that's lasted 20 years?"

He said, "Well, I hope my marriage does." But Lotus lasted 13 years before it was bought and made into a brand inside some other corporation. So, 13 years was sort of a major software company. 20 years for a non-profit, Electronic Frontier Foundation – there's

something going on, that there's the long-term, stable – even in this economic downturn – all of these organisations, on the board of a bunch of them, are all doing very well. We're all growing along. There's a need for it.

So, this high-tech non-profit as building infrastructure, I think, is a very important, new capability, and I think we'll see it grow. And organisations that know how to deal with the non-profits, whether they're a government or corporations, and leverage this new way of harnessing creative, innovative, participatory, transparent, open organisations, will be a critical form of whether it's a successful country, successful companies.

43:28 .COM AND .ORG INTERACTION

One of the most interesting things I've seen out of the Internet, the growth of the Internet, is something that probably most people don't really see. It's the ontology. It's dividing things – the world – into dot-gov, dot-com, dot-edu, dot-mil, and dot-org. It's fascinating, right, to go and cut up organisations of people into those top-level ways of viewing the world. It was done by Jon Postel, and I think it's one of the most insightful things of the Internet. So, what we've seen come about is this interaction between the dot-com and the dot-org that it's, I think, potentially one of the most hopeful things to come out of this whole thing.

So, the gasoline of capitalism, right, you can get people really excited in money and power going towards making something happen. But at some point, it hits a level of maturity, and does it end up then into a monopoly organisation of a corporation that owns it, or did it actually diffuse out into the dot-org world, the non-profit? Does it become all of ours? And sometimes seeing one, sometimes seeing the other. I'm

hoping that, for instance, in the books area – as books are going online, it could go so that it was really spearheaded and the power was put towards it with the dot-com, but then they belong to all of us once again. So, we had – a lot of the early development in book scanning was done with government funding, but really it hit stride when Google said, "We're gonna just go and digitise all of it." And that caused Yahoo! and Microsoft to go and say, "Oh, well, we better do something about this," and they invested in an open organisation. The Internet Archive – we worked with lots of libraries. Microsoft put \$10 million towards getting the deployment of book scanning out there, and this came after foundations had put in a couple million dollars into the R&D.

So, R&D was the foundations. The VC – Microsoft put a power stroke, but they didn't want to own it. They just wanted somebody else not to own it – interesting. So, Google is powering through it. They're working very hard. But we might see – and I'm hoping that it does happen this way – that the enthusiasm and power of a Google to go and say, "We can build a future that has all the books online," doesn't mean that all the books are owned and controlled by Google at the end of the day, that they belong to all of us, that the technology – and more important probably, the idea that it's possible – is brought to us by a corporation. And now, it's time for the public sector to stand up to it.

So, we now have France going and investing 1 billion euros to go and digitise their heritage. So, that seems like a large number, but it's actually very small in the course of how much they spend on their libraries. If each country were to go and do this, we would have – or even one-tenth of that – we would have all of the world's libraries available online within the public sphere. So, where we have dot-coms

starting to really do some powerful things, it might be that the dot-orgs, the non-profit, the more distributed Internet, that is a possibility, is how it actually is rendered in the future, so there isn't centralised control. Will it happen? Don't know, but I think it's our opportunity.

47:00 QUESTION FOR THE ORACLE

If I were to go to the Oracle at Delphi now and try to ask some questions to try to get some idea of what it is going to happen and what it is we should do now, the question I'd really try to – and I do – try to understand is what can we do now within our own small worlds to build a future we want to live in? What are the tipping points that are going on now? If we were to go and work hard at something to go and make it so that the future really is within whatever view of the world it should be – my world is I want an open information environment that's creative and fun to be in. I want whacko people to be able to go and share information in wild and interesting ways. I want to be surprised and have things come about.

Anytime I'm in a different location or a different community, I want them to be very different from each other. I want a diversity. I want language to be flowering. I want poetry and all sorts of different kinds of things. This is the world I want. This is not necessarily what other people want. So, if this is my utopia, what should I be working on now to create that, and what are the pivot points? What are the things that are – if you just push that asteroid a couple degrees now, the world's very different later. And what actually doesn't matter? What is just redundant with what other people are doing? What's just playing out the roles, or being part of the cog in the wheel, that frankly it doesn't really matter if you just do another one of those – eh. But what are the important things now that are going to shape the future?

48:56 OPEN INFORMATION

The easiest pieces that I see right now that if we were to invest a small amount of time and effort, societally, if we were to just build an underlying infrastructure of a certain set of services that are free as air, that are just like government – what's the equivalent of the road system or clean water, but in the information age? And how do we go and organise ourselves to create those? I think it's open source software. I think it's things like Creative Commons licenses – really smart on Larry Lessig's part. What we need is a good set of licenses. It cost a couple million dollars to go and build those, and it's an organisation that costs about two million dollars a year. And we now have licenses that allow us to share. Everybody benefits because Larry Lessig and a small band of people set up a set of licenses. It's what government should have been doing – it wasn't. It's going on in non-profits now.

But what are some more of those that are high leverage point parts of infrastructure? And it turns out that a lot of these don't cost very much. Some of it's regulatory, going and making sure that there are open bandwidth so that there's Net Neutrality, so that there's possibility of new entrance, so that there's swathes of the wireless spectrum that if you just use any – if you use WiFi, you're allowed to use it.

WiFi came about in this awful little part of the spectrum that was where microwave ovens were, and that was the only reason why it wasn't licensed already. It wasn't packaged and sold to some corporation, and we have an amazing future because there was a piece of unlicensed spectrum. What's the equivalent now in the content area? What's the information that wants to be free? Is it government information with no copyrights, universally accessible? That's available

in the United States, or at least, that's the law in the United States, but is it really happening or is it getting locked up? Is Europe going with open information?

Regulatory information is not even free. Laws are not even open and free. I think there's an underlying bedrock of information that we need to operate as a society. It doesn't cost us very much to put up there and put out there. It doesn't require selling it off to an organisation, single corporation to own and control. I think small investments like that can help us create a fantastic and interesting world going future.

51:34 BUSINESS OPPORTUNITIES IN CREATING OPEN INFORMATION

There's this myth that openness dissolves and is anti-capitalist, and that's absolutely flip opposite. If you take capitalism, in its best sense, as a competitive environment where lots of people get to contend and play in, openness is absolutely critical. So, if we take WiFi as an example, where it's a very narrow piece of spectrum where we've had lots of interesting things happen, billions of dollars of organisations have been able to be created. Lots of interesting, unpredictable things happen, but you take a couple megahertz up, couple megahertz down, you have owned and controlled and it's a desert.

So, in a very narrow – it's because it was open that we could do something interesting. We're seeing the same thing with software. The Internet, because it was – the protocols were open – it caused billion dollar corporations to come about. So, openness is the best thing for capitalism. Closed is the best thing for monopolies, and monopolies are a starved desert of information and innovation.

52:44 THINKING MACHINES: COMPUTERS AS PARTNERS

So, in the 1980s, there was this idea that you could build a question answering machine. You can make a machine that's interesting to hang out with, a machine that you wouldn't instruct. It would be fun to be with. It would know enough, and the computing has been going along on the right curve. The storage of knowledge has been absolutely on the right curve. And we're now starting to see the interaction between the computing and the knowledge bear fruit in ways that are somewhat similar to what it is was dreamed of back then, where you could say, "Computer, please tell me the answer to this."

We're starting to see the first cell phone apps that can answer questions that you ask just that way, and the expectation of the computer in Star Trek or in 2001: A Space Odyssey is starting to come about. It's so slow and so subtle that it's hard to see it arrive at any particular birthday. They're not shocking and being announced all at once, but the search engines, I think, were one of the first biggies that really blew people's mind, where you could ask it a question. And it didn't really know the answer. It didn't really put together the answer. It just regurgitated an answer that somebody else came up with.

Wikipedia really is basically that. I mean, you can almost type the titles of articles, and I know people that do. They don't search Wikipedia, because articles are designed around pretty much all the questions and answers that you might want to answer. And people put it in, and it's the idea of going and crowd sourcing, or everyone putting together that knowledge base was not what we were expecting 20, 30 years in the artificial intelligence labs around the world.

But it was a pretty interesting answer, but we're starting to see computers start to put facts together from multiple sources, and be able to make those available to us in interesting and new ways. That, I think, we're going to become absolutely dependent on, so that it's going to be – the idea of being offline is going to make us sort of really be on vacation, or almost asleep. The idea of being very far away from our personal, digital assistants, it's going to be really difficult. I mean, we're going to be really almost going into a meditation or a trance that's separated from the world.

I wonder – if you were to take a bank now, and you're going to have the choice, the draconian choice, of kill all the people or kill all the computer systems in the bank, and you were trying to maximize the stock value of the thing that's left over with, I think you'd probably get rid of all the people. The computers – and what's known by those computers – is much more important than what's actually known by the people. The people are more replaceable than the computers. That's a little odd.

What does it mean if we do that with another level out? So, we're dependent on question-answering capabilities of our devices, of our car navigation systems, of the Internet search engines, whatever they become. And when that really evolves – education system is hopefully going to change enough. But it means that these computers are weaving into our lives in a way that we can look to the science fiction of 50 years ago and see it absolutely coming true. And I find it very exciting, very interesting. I'm finding it fun to play with my computer. They're interesting things you can play with, and I learn from my computer, and is it teaching me? In some fairly real way, yes. I see this area is going more and more in this direction as computing, the

content, and the networks all come together to build a library that's worth talking to.

57:15

A NEW INFRASTRUCTURE OF QUESTION AND ANSWERING

So, the search engines of today, where you go and ask a question and you get a whole document that's kind of an answer, is very library-like. When I went to library school, we were taught a reference interaction. So, when a patron comes and asks a question, you don't just go and say, "1943." You say, "Here's an answer toward – to that," so that people can explore not just the answer, but around it, what's the context, who said so, that type of thing.

That's what the search engines are. People's tolerance for this is low. They would really much rather have, "1943," or they'd like to know, "I should be doing this." And the evolution of those sorts of systems is starting to come about, that's much more general than going to the United Airlines site to go and find out whether your flight's delayed, that we're starting to see question-answering machines come about, whether it's Siri or Wolfram Alpha. They're really very sophisticated and so, people are going to want these much more quick interactions.

So, I think we'll see the evolution that you'll use libraries and search engines kind of like libraries in the future, where that's when you're kind of exploring. You're not quite sure what. You want a more in-depth answer. You're not quite sure who to trust, but much more of our interactions with our technology is going to become much more fluid towards natural language and expecting precise answers coming back. Very exciting. And there are going to be people that try to –

organisations – that try to extract factual information from the fluid search engine-like stuff to go and build these crisply-interactive technologies.

And we'll see, I think, some of the fruition of these systems and start to be pretty interesting and fun. I'd like to see how these things interact with each other, because right now they're in silos, these silos being different organisations where – can we go and build a World Wide Web of question-answering machines? We've built a World Wide Web of search engines and web-type things, but can we go and make it so that – I know how to answer questions about baseball better than anybody. So, can we make one of these answer engines go and probe me when there's an appropriate time, so that there's a new infrastructure of question answering, as opposed to just document shovelling?

Open question, we haven't even seen people try. Mostly they're just trying to get anything to work at all, and some of the early things like Wolfram Alpha are extremely exciting to me.

1:00:02 INTERNET IS ABOUT THE PEOPLE

Any place there are people will play a role in the future of the Internet, because fundamentally the Internet really is about the people. Now, they may go and give control to other organisations that'll Hoover them all together but I think, all in all, really this thing comes from us. And if we let it get out of control, if we have governments that are asleep, we can end up in some pretty bad times of focused, monopolistic control. But it's so inefficient to run an economy, based on monopolies that hopefully it will eventually dissolve. One of the best things is to just not have them happen in the first place, but will Europe play a role? Absolutely, as long as they're sort of alive and

kicking, and we'll see India and China play – and microclimates of people that are distributed over the world will play a role, as long as we keep the Internet flat for new applications to be created and distributed without approval from somebody that controls either a pipe or a device.